

**Kovács Kristóf – Pléh Csaba**

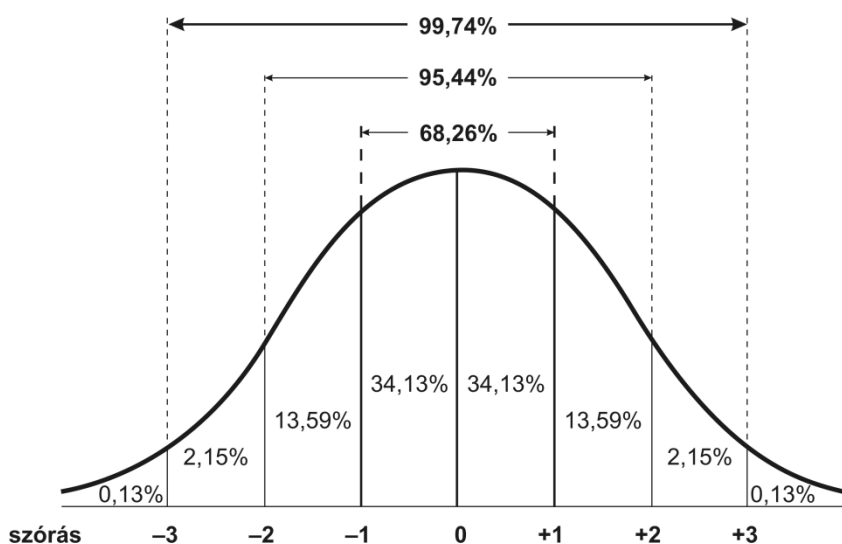
Eszterházy Károly Főiskola

[kristof340@gmail.com](mailto:kristof340@gmail.com)[pleh.csaba@ektf.hu](mailto:pleh.csaba@ektf.hu)**Han van der Maas**

Universiteit van Amsterdam

[h.l.j.vandermaas@uva.nl](mailto:h.l.j.vandermaas@uva.nl)**AZ ADAPTÍV TESZTELÉS LEHETŐSÉGEI AZ IKT-ALAPÚ OKTATÁSBAN****1. Egyéni különbségek és a „két szórás probléma”**

A gyerekek különböző személyiséggel, érdeklődéssel, és motivációval érkeznek az osztályba, ezek az egyéni különbségek pedig az egyes tanulók esetében meghatározzák a haladás lehetséges ütemét. A legtöbb pszichológiai vonás, és különösképp a kognitív képességek, normális vagy haranggörbe-eloszlást követnek (1. ábra). Ez azt jelenti, hogy az egyes tantárgyak elsajátításához szükséges készségek tekintetében a gyermekek nagyjából kétharmada esik az átlagtól egy szórásra belülre: a képességek pszichológiája szempontjából őket tekinthetjük „átlagos” diáknak.

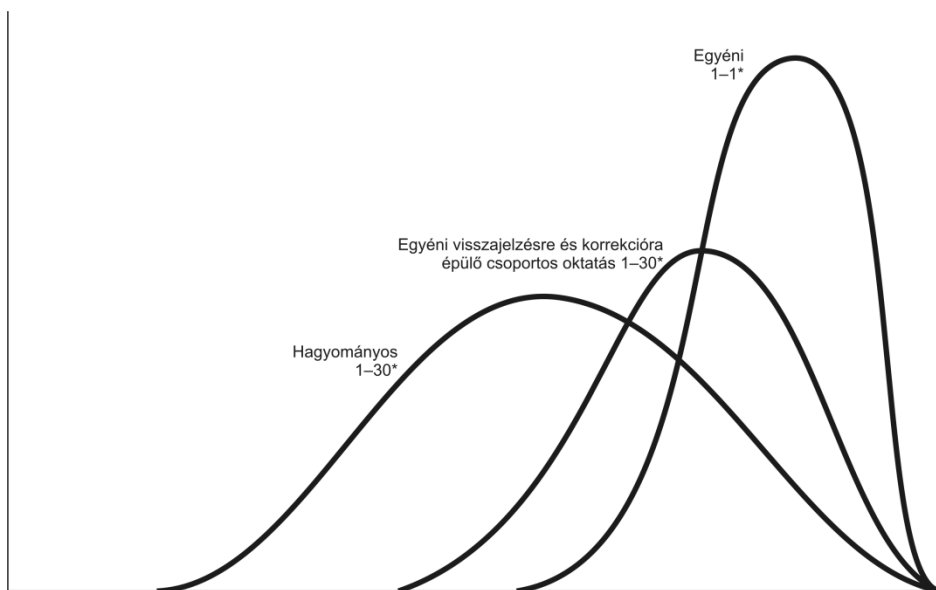


1. ábra: A mentális képességek normális eloszlása az egyes szórások közti, valamint az összesített gyakorisági értékekkel

A hagyományos tantermi oktatás során, vagyis amikor egyetlen tanár tanít nagyjából 30 gyereket, az eloszlás két szélén lévő diákok lesznek azok, akik számottevően eltérnek

a többségtől a tekintetben, hogy mennyire gyorsan tudnak haladni, gyakorlás során pedig mennyire nehéz feladatokkal képesek megbirkózni. A letehetősebbeket az unalom fenyegeti, a tempó számukra túl lassú, a gyakorlatokkal gyorsabban végeznek, mint a többiek, és ez folyamatos feladatokat ró a tanárookra. A leglassabbak viszont nem képesek lépést tartani a többiekkel, kudarc-élmények érik őket, ezért csökken a motivációjuk. Mindennek következtében sok tanuló nem képes maradéktalanul kiaknázni a képességeit, és eljutni arra a tudásszintre, amelyre elvben eljuthatna.

Számos osztálytermi kutatás mutatott rá arra, hogy a tutori oktatás, amikor minden gyermekkel külön foglalkozik egy tanár, sokkal hatékonyabb, mint a csoportos tanítás. A különbség mértéke mellbevágó: kontrollált vizsgálatok szerint (Bloom, 1984) az egyéni oktatásban részesülők két szórásnyival jobb teljesítményt érnek el, mint a csoportos oktatásban résztvevők (2. ábra). Ez azt jelenti, hogy a tutori oktatásban részesülő diákok átlagos teljesítménye jobb volt, mint a hagyományos oktatásban részesülő gyerekek 98%-áé. A vizsgálatok során a gyerekeket véletlenszerűen sorolták az egyes csoportokba, és a csoportok között nem volt különbség az előzetes tudás és motiváció átlagos szintje között.



2. ábra: (Bloom, 1984, 5. o. alapján): A különféle oktatási módszerek eredményességének összehasonlítása. \*Tanár-diák arány.

A vizsgálatok során az egyéni oktatás mellett kipróbáltak egy olyan elrendezést is, amely tartalmaz visszajelzést és korrekciót – sztenderdizált tesztek segítségével –, ugyanakkor az oktatás a hagyományos, csoportos formában zajlik. Ez a megoldás is számottevően felülmúlta a hagyományos oktatással elért eredményeket, azonban meg sem közelítette a tutori oktatás hatékonyságát: a gyerekek itt átlagosan egy szórásnyival teljesítettek jobban, vagyis az átlagos gyerek teljesítménye jobb volt, mint a konvencionálisan oktatott gyerekek 84%-ának a teljesítménye. Érdeemes megjegyezni,

hogy ennek a fajta oktatásnak a legfőbb erénye, az egyéni visszajelzés és korrekció, a tutori tanításnál eleve nem központi jelentőségű, mivel a tapasztalatok azt mutatták, hogy szinte egyáltalán nincs szükség utólagos korrekcióra: a tipushibák az egyéni oktatás jellegénél fogva már a tanítás során felszínre kerülnek, és azonnal orvosolhatók.

A Bloom (1984) által összefoglalt vizsgálatok számos további fontos eredményre is rámutattak. Az egyik, hogy az alternatív oktatási módszerek számottevően csökkentik a csoporton belüli variációt, mint ahogy ez a 2. ábrán is látható. A tutori oktatásban részesülők 90%-a, a visszajelzéssel és korrekcióval kiegészített csoportos oktatás alanyainak pedig 70%-a esett abba a tartományba, amelybe a hagyományos oktatásban részesülő gyerekek legjobb 20%-a tartozott.

A második, hogy az egyéni oktatás közvetlen hatékonysága is kiemelkedően magas volt: a gyerekek másfélszer annyi időt töltöttek a feladatokkal, mint a hagyományos csoportban. Végül pedig számottevő különbség mutatkozott a gyerekek attitűdjében is: a tutori oktatásban részt vevők nyilatkoztak legkedvezőbben a tanulásról és ők voltak a legérdeklődőbbek, a hagyományos oktatásban részt vevők pedig a legkevésbé. Úgy tűnik tehát, hogy az egyéni oktatás nem csak eredményesebb, hanem élvezetesebb is. Továbbá a drasztikusan javuló teljesítmény a lecsökkent variációval együtt azt mutatja, hogy tutori típusú oktatás segítségével a diákok közti előzetes különbségek nagymértékben csökkenthetők, és szinte mindenki jóval magasabb szintű teljesítményre képes, mint egy iskolai osztály átlaga.

Az egyéni oktatás természeténél fogva alkalmas arra, hogy a tanárok alkalmazkodjanak a tanulók képességeinek egyéni különbségeihez. Nyilvánvalóan nem megoldható azonban, hogy a hivatalos iskoláztatás során minden egyes gyerekre jusson egy, csak vele foglalkozó tanár: ez nem csak túl költséges, de praktikusán sem megvalósítható. Ezért a kérdés az, hogyan lehet olyan módszereket találni, amelyek hatékonyság tekintetében felveszik a versenyt a tutori típusú tanulással, azonban életszerűen összeegyeztethetők a csoportos, osztálytermi oktatással. Bloom ezt nevezte „két szórás problémának”.

## **2. IKT az oktatásban és a számítógépes adaptív tesztelés**

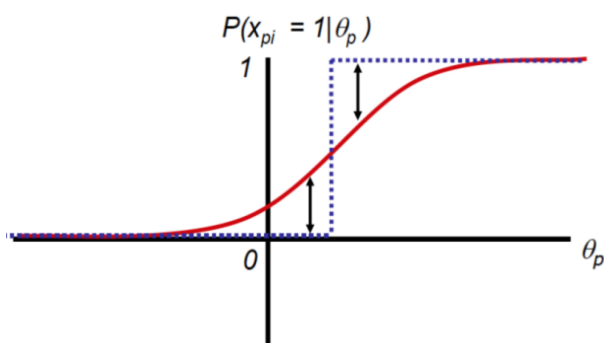
A két szórás probléma lehetséges megoldása a „számítógépes kognitív tutorálás” (Corbett, 2001), illetve általában az IKT-eszközök kiegészítő használata az osztálytermi oktatás mellett. A modern technológia lehetővé teszi, hogy a tanulók tutor nélkül is a saját képességeik szerint gyakorolhassanak, vagyis mindenki olyan nehézségű feladatot oldjon meg, amely számára még éppen megoldható, és így mindenki maximálisan használja a képességeit. Ezáltal a legtehetségesebb gyerekek sem unatkoznak, a leglassabbaknak pedig nem okoz frusztrációt a tanulás: azáltal, hogy a saját tempójuk szerint haladhatnak, nem veszítik el a motivációjukat.

Ahhoz, hogy tisztában legyünk az egyes tanulók képességszintjével, pszichometriai mérést kell végeznünk. Ha pedig az oktatás amúgy is felöleli IKT-alapú eszközök használatát, érdemes a mérést is ennek segítségével, nem pedig papír-ceruza eszközökkel végezni. A képességek egyéni különbségeinek IKT-alapú mérése a számítógépes adaptív tesztelés körébe tartozik, amely pedig elválaszthatatlan a pszichometriai fejlődésétől, vagyis a modern tesztelmélet (item-válasz elmélet) megszületésétől.

A korábban domináns klasszikus tesztelmélet lényege, hogy az egyéni eredményt egy normacsoport összpontszám-eloszlásához viszonyítják. Ez azt jelenti, hogy teszten elért nyerspontszám önmagában nem jelent sokat, amíg össze nem hasonlítják egy adott minta korábban, a sztenderdizálás során mért eredményével. Vagyis először egy reprezentatív minta tölti ki a tesztet, majd kiszámításra kerül a pontszámaik átlaga és szórása, hogy később ezekhez tudjuk viszonyítani az egyedi eredményeket. Mivel a mentális képességekre az 1. ábrán látható normális eloszlás jellemző, a szórásban kifejezett érték (Z-érték) alapján megmondhatjuk, hogy az adott eredmény a népesség hány százalékaénak a pontszámánál jobb: ezt fejezi ki az úgynevezett centilis érték.

A klasszikus tesztelmélet megfelelő elméleti keretet nyújt a pszichometriai méréshez, azonban számos hátulütője akad. Először is: noha a tesztek egésze rendelkezik megbízhatósági (reliabilitási) mutatóval, amely jelöli az egyes tesztek mérési pontosságát, az egyes eredményeknek nincs hibataromány, így az egyes konkrét tesztteredmények pontossága sem megbecsülhető. Másrészt csak a teszt egésze képes mérni, hiszen az összpontszámot hasonlítjuk a normacsoportéhoz, ennél fogva az egyes teszt-ítemek nem felcserélhetők, a teszt pedig nem bővíthető, csak teljes újrasztenderdizálással együtt. Ez azt jelenti, hogy az egyes feladatok önmagukban semmilyen információt sem nyújtanak. Harmadrészt, a teszt használhatósága nagyban függ a normához használt mintától, amelynek a normális eloszlás minden szegmensét arányosan le kell fednie. Végül pedig, mivel a tesztek hosszúsága, és így az ítemek száma korlátozott, a tesztkészítés során az átlagos nehézségű feladatokra kerül a hangsúly. Készíthető persze a normális eloszlás alsó vagy felső részét megcélzó teszt, ez azonban szintén szűk tartományú lesz, és éppen az átlag körüli képességek megkülönböztetésére nem lesz alkalmas. Vagyis a klasszikus tesztelmélet alapján készült tesztek mérési tartománya szükségszerűen szűkös.

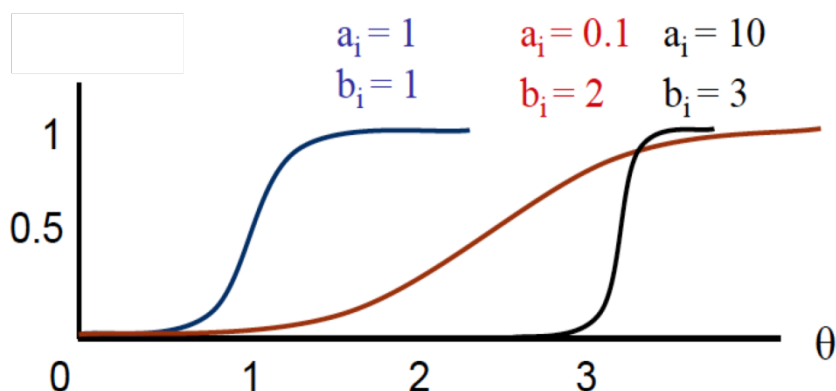
A klasszikust meghaladó modern tesztelmélet, vagy ítem-válasz elmélet (ld. Hambleton, Swaminathan, & Rogers, 1991) alapfeltételezése, hogy a jobb képességűek nagyobb valószínűséggel válaszolnak helyesen egy adott kérdésre, mint a rosszabb képességűek, függetlenül bármilyen más jellemzőjüktől. A modern tesztelmélet vagy ítem-válasz elmélet tehát probabilisztikus: az egyes feladatok saját úgynevezett ítem-paraméterekkel rendelkeznek, amelyek megjósolják, hogy a mérendő képesség egy adott szintjén mekkora valószínűséggel oldják meg az adott feladatot.



3. ábra: egy példa az itemjelleg-görbére (pirossal), amely egy adott feladat helyes megoldásának valószínűségét mutatja a képességszint ( $\theta$ ) függvényében. A feladat nehézségét a szaggatott vonal jelzi.

Így minden egyes item esetében külön meghatározható a helyes válasz valószínűsége a képességszint függvényében. Ezt fejezi ki az úgynevezett itemjelleg-görbe. A kétparaméteres modellben (Birnbbaum, 1968) minden egyes itemnek két jellemzője van: a nehézség és a diszkrimináció. A nehézség-paraméter („b”) a képesség-skála azon értékének felel meg, amelynél a helyes válasz valószínűsége 0.5. (3. ábra). Minél magasabb ennek a paraméternek az értéke, annál magasabb képességszint kell ahhoz, hogy 50%-os valószínűséggel oldjuk meg a feladatot.

A nehézségen kívül a feladatok másik általános tulajdonsága a diszkriminációs paraméter („a”), amely az itemjelleg-görbe meredekségével arányos (4. ábra). A diszkriminációs paraméter segítségével megállapítható, hogy az adott item mennyire differenciál, vagyis mennyivel nagyobb valószínűséggel válaszol helyesen egy magas képességszintű egyén, mint egy alacsony képességszintű. Minél meredekebb az itemjelleg-görbe, annál inkább alkalmas a feladat arra, hogy a helyes vagy helytelen válasz alapján megkülönböztessük az alacsony és magasabb képességűeket. A magas diszkriminációjú itemek tehát több információt szolgáltatnak. Az alacsony diszkriminációjú feladatok viszont rossz mérési tulajdonságokkal rendelkeznek, mivel a helyes-helytelen válaszból csak nagyon pontatlanul tudunk következtetni a mögöttes képességszintre.



4. ábra: Három, különböző nehézség- és diszkriminációs paraméterrel rendelkező feladat itemjelleg-görbéje

A modern tesztelmélet számos előnnyel rendelkezik a klasszikussal szemben. Egyrészt az egyes feladatokkal is lehet mérni, nem csak a teljes teszttel. Ebből következően nem minden válasz „ér ugyanannyit”, szemben a klasszikus tesztekkel. Másodsor: a mérés pontossága minden egyes eredmény esetében megbecsülhető. Harmadsor: a paraméterek mintafüggetlenek, azok számítása több részminta segítségével is lehetséges, amelyek együtt sem kell, hogy lefedjék az összes képességtartományt. Végül: az item-válasz elmélet lehetővé teszi az adaptív tesztelést.

Valójában a számítógépes adaptív tesztelés (van der Linden & Glas, 2002; Weiner & Dorans, 2000) teljes egészében a modern tesztelméletre épül. A számítógépes adaptív tesztelés (CAT) során, a papír-ceruza tesztekkel szemben, nem egy kész feladatsort használnak, hanem egy úgynevezett item bankot, amely akár több száz feladattól is állhat és amelyből az adaptív algoritmus válogat. A teszt egyénekre szabása úgy valósul meg, hogy az algoritmus a kitöltő becsült képességszintje alapján választja ki az egyes itemeket, azok diszkriminációs képessége és nehézsége alapján. Vagyis az adaptív algoritmus a következő kérdést mindig az az előző kérdésekre adott válaszok alapján adja.

A folyamat a gyakorlatban úgy valósul meg, hogy elindul a teszt egy adott képességszintről, ami általában az átlagos képességszintet jelenti. Az algoritmus kiválaszt egy feladatot az item bankból, amelynek a nehézsége közel áll ehhez a szinthez. Regisztrálja, hogy a beérkező választ helyes-e, majd ennek megfelelően ad egy úgynevezett likelihood-beclést a képességszintről, valamint megadja ennek a beclésnek a hibáját.

Több feladatra adott válasz után a helyes és helytelen válaszok sorozata adja meg a kitöltő eredményét: az algoritmus pedig minden egyes új item után megbecsüli, hogy az adott nehézségű feladatokra érkezett helyes illetve helytelen válaszok együttes előfordulása mely képességszinten a legvalószínűbb. Vagyis lényegében fordított valószínűséget számol az item-paraméterek alapján: abból, hogy az egyes feladatokat különböző képességszinten milyen valószínűséggel oldják meg, megadja a

legvalószínűbb képességszintet, amely mellett a helyes és helytelen válaszok empirikus mintázata előfordulhatott.

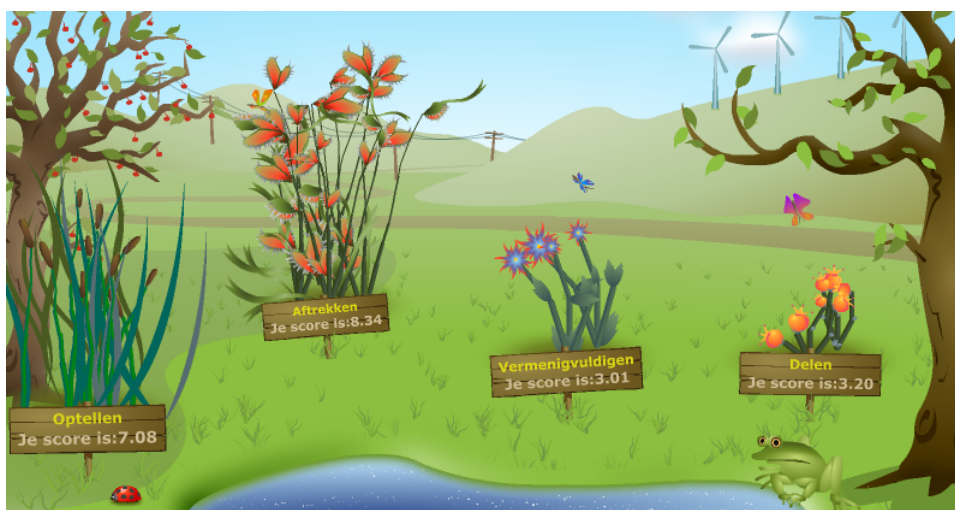
Minél jobban diszkriminál egy feladat, annál inkább csökken a hibatartomány a feladat megoldását követően. A becsült képességszint alapján az algoritmus kiválasztja a következő itemet, a leggyakrabban használt eljárás során minden esetben az adott szinten a legtöbb információt szolgáltató item lesz a következő kérdés. Az itemekre adott válaszokkal tehát egyre inkább csökken az esetleges tévedés mértéke is. Az eljárás rendszerint addig tart, amíg a becsült érték hibatartománya le nem csökken egy előzetesen meghatározott szint alá, de megszabhatunk egy maximális kérdésszámot is, vagy egy képességszintet, amely alatt vagy felett a teszt már nem mér.

A CAT a gyakorlatban tipikusan azt jelenti, hogy hibázás után könnyebb, helyes válasz után nehezebb feladat következik. A CAT lényege, hogy míg a hagyományos tesztben a kitöltő számára túl könnyű és túl nehéz feladatokkal is találkozhat, az adaptív teszteset alkalmazkodik a képességszintjéhez, így a tesztre szánt idő legnagyobb részében a saját képességszintjének megfelelő feladatokat kap. Ez egyrészt felgyorsítja a teszteseti folyamatot, másrészt növeli a mérés pontosságát, és megkíméli a fölösleges frusztrációtól. A CAT további előnyei közé tartozik, hogy nincs kész – tehát ellopható és betanulható – megoldókulcs, mert minden teszt „személyre szabott”, másrészt az item bank cserélhető és bővíthető.

### **3. Rekentuin (Matekliget): egy adaptív online matematika-gyakorló rendszer**

A CAT alapelvei kiterjeszthetők az oktatás világára is, és általuk lehetővé válik a személyre szabott tanulás. Egy adaptív rendszer mindig a felhasználó aktuális teljesítményének megfelelően válogatja a feladatokat, úgy, hogy azok mindig kihívást jelentsenek, de ne legyenek túl nehezek. Így a tanulók a saját képességeik szerint gyakorolhatnak és a saját egyéni ütemük szerint fejlődhetnek.

Az Amszterdami Egyetem pszichológusai által kifejlesztett Rekentuin (Math Garden, magyarul nagyjából: Matekliget) pontosan ilyen adaptív gyakorló környezet: a számítógépes adaptív tesztesetet ötvözi az egyénre szabott gyakorlással. A rendszer úgy van beállítva, hogy a gyerekek a kapott feladatok többségét várhatóan képesek legyenek sikeresen megoldani. Egy nagyon jó matekos, aki a könnyű kérdésekre kivétel nélkül helyes választ ad, automatikusan egyre nehezebb kérdéseket kap, miközben a gyengébb matekosok is a saját szintjükön méretnek meg. Ezért a gyerekek számára a Rekentuin használata sikerélmény, nem pedig kudarc forrása. Ugyanakkor a feladatok egyúttal számítógépes adaptív tesztként is működnek, amely alapján a rendszer minden egyes válasz után újra becsüli a felhasználó képességszintjét.



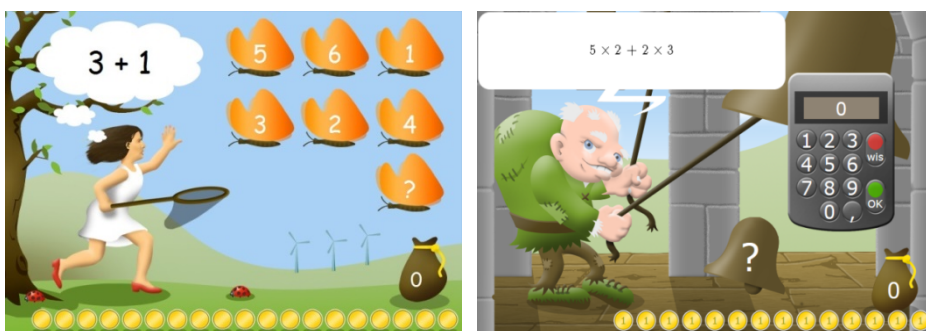
5. ábra: a Rekentuin főoldala az alapműveletekkel és a hozzájuk tartozó növényekkel

A Rekentuin négy éves kortól használható (Hollandiában ez a beiskolázás életkora), a nehezebb feladatok pedig még középiskolások számára is kihívást jelenthetnek. A matematika gyakorlása játékos környezetre épül: az egyes feladat-típusok (összeadás, kivonás, szorzás, osztás, számolás, órák stb.) egy kert növényeinek felelnek meg: a felhasználó a feladatok megoldásával tudja gondozni a kertjét, amely annál szebb lesz, minél többet gyakorol (5. ábra). Ha viszont ezt elmulasztja, a növények hervadni kezdenek. A Rekentuin a rendszeres gyakorlást és a fejlődést egyaránt jutalmazza.

Azon túl, hogy ötvözi az IKT-alapú matematika-oktatást és a számítógépes adaptív tesztelést, a Rekentuin két további új megközelítést is alkalmaz. Az egyik az, hogy adaptív rendszere nem csak azt veszi figyelembe, hogy egy problémát helyesen oldanak-e meg, hanem a matektudás részének tekinti azt is, hogy a gyerek mennyire gyorsan jut el a megoldásig. Mivel a számítógépek könnyen rögzíteni tudják a válaszadás idejét, ezt az információt be lehet építeni az értékelésbe, így a megoldáshoz szükséges idő járulékos információt szolgáltat a gyerek matektudásáról. Ezt pedig a rendszer figyelembe veszi a pontozásnál, még hozzá a hibás válaszok esetében éppúgy, mint a helyeseknél, ezáltal ügyelve arra, hogy hosszabb távon ne díjazza a tippeléssel kapott válaszokat.

A gyerekek érdemeket gyűjthetnek a helyes válasszal, annál többet, minél gyorsabban válaszolnak – de veszíthetnek is érdemet, ha hibásan válaszolnak. A játékok tehát arra ösztönzik a gyerekeket, hogy olyan stratégiát válasszanak, amelynek segítségével gyorsan és pontosan tudnak válaszolni. Mivel mindenki a saját szintjének megfelelő feladatokat kap, ugyanannyi játékkal minden gyerek ugyanannyi érdmet gyűjthet össze, a képességszintjétől függetlenül. A válaszadás idejének mérése azzal az előnnyel is jár, hogy a gyerekek kaphatnak könnyű feladatokat, mert így is lehet hatékonyan tesztelni. Vagyis kaphatnak olyan problémákat is, amelyet akár 75%-uk is helyesen megold, ez pedig nagyon erősíti a tanulók motivációját.





6. ábra: példák a Rekentuin feladataira

A másik újítás az, hogy sztenderd kétparaméteres modellel szemben a készítő az adaptív számítógépes tesztelésre épülő képességszint-beccslés során a sakkban a játékerő megállapításakor használt, Élő Árpád magyar származású fizikus által kidolgozott Élő-pontszámhoz hasonló számítást alkalmaznak (Maris & van der Maas, 2012). Vagyis ahelyett, hogy előzetesen egy mintán megbecsülnénk az item-paramétereket, a Rekentuinben minden feladat átlagos nehézségről indul, a gyerekek pedig átlagos képességről. Ezt követően a program az Élő-pontszámhoz hasonló algoritmussal módosítja ezeket. Tehát gyakorlatilag a gyerekek és a matekfeladatok ellenfelek, és mindkettőnek megvan a saját „játékereje”. Ez a gyerekek esetében a matektudásukat fejezi ki, a problémák esetében pedig a nehézséget tükrözi. Ha a gyerek helyesen válaszol, akkor ő „nyert”, és a pontszáma valamelyest növekszik. A feladat viszont „vesztett”, a pontszáma pedig csökken.

A Rekentuin tehát nem egy CAT algoritmus, hanem egy fejlesztő-utánkövető rendszer, mivel a feladatmegoldás itt valójában nem teszt, hanem gyakorlás. A gyakorlás során keletkező valamennyi adat eltárolódik, a tanárok pedig ezek segítségével nyomon követhetik a diákjaik egyéni fejlődését. A központi adatbázisnak köszönhetően az egyes eredmények különböző referencia-csoportokhoz is hasonlíthatók, így például megtudható, hogy egy diák matektudása elmarad-e a kortársaitól, vagy épp ellenkezőleg, gyorsabban halad, mint a többiek. Vagyis a tanárok vagy a szülők pedig részletes, profil-alapú információkat nyerhetnek a rendszert használó gyerekek fejlődéséről: a Rekentuin képes visszajelzést adni az egyes gyerekek vagy akár csoportok (pl. osztályok) teljesítményéről és fejlődéséről, össze tudja hasonlítani azt egy adott csoport (pl. egy iskolai osztály) átlagos teljesítményével és fejlődésével, továbbá hibaelemzés segítségével részletes profilt készít az egyes erősségekről és meghatározza a fejlesztendő területeket.

A Rekentuin teljesen webes alapú: használatához nincs szükség telepítésre, csak egy számítógépre vagy tabletre és internetkapcsolatra. Hollandiában, Belgiumban, és az Egyesült Királyságban jelenleg közel 1200 iskolában összesen 120 000-nél is több gyerek használja a Rekentuin-t, vagy annak angol nyelvű változatát, a Math Garden-t. A rendszer 2009-es indulása óta eltelt öt évben több mint 300 millió egyedi válasz érkezett az egyes feladatokra, ma már naponta 750 000 körüli választ rögzít a rendszer. Ezek alapján az egyes feladatok nehézsége rendkívül pontosan kalibrálható.

A kutatás a TÁMOP-4.2.4.A/2-11/1-2012-0001 Nemzeti Kiválóság Program című kiemelt projekt keretében zajlott. A projekt az Európai Unió támogatásával, az Európai Szociális Alap társfinanszírozásával valósul meg.

### Hivatkozások

- Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In F. M. Lord & M. R. Novick (Eds.), *Statistical Theories of Mental Test Scores* (pp. 395–479). Reading, MA: Addison-Wesley. Retrieved from <http://ci.nii.ac.jp/naid/10011856529/en/>
- Bloom, B. S. (1984). The 2 Sigma Problem: The Search for Methods of Group Instruction as Effective as One-to-One Tutoring. *Educational Researcher*, 13(6), 4–16.
- Corbett, A. (2001). Cognitive computer tutors: Solving the two-sigma problem. In *User Modeling 2001* (pp. 137–147). Berlin Heidelberg: Springer.
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of Item Response Theory*. London: Sage Publications.
- Maris, G., & van der Maas, H. (2012). Speed-Accuracy Response Models: Scoring Rules based on Response Time and Accuracy. *Psychometrika*, 77(4), 615–633. doi:10.1007/s11336-012-9288-y
- Van der Linden, W. J., & Glas, G. A. W. (2002). *Computerized Adaptive Testing: Theory and Practice*. New York: Kluwer Academic Publishers.
- Weiner, H., & Dorans, N. J. (2000). *Computerized Adaptive Testing: A Primer* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.